Crystal structure of rat α-parvalbumin at 1.05 Å resolution

CHRISTOPHER A. BOTTOMS,[1] JONATHAN P. SCHUERMANN,[1] SAYEH AGAH,[2] MICHAEL T. HENZL,[2] AND JOHN J. TANNER[1,2]

[1]Department of Chemistry and [2]Department of Biochemistry, University of Missouri–Columbia, Columbia, Missouri 65211, USA

(RECEIVED December 18, 2003; FINAL REVISION April 1, 2004; ACCEPTED April 1, 2004)

Abstract

The crystal structure of rat α-parvalbumin has been determined at 1.05 Å resolution, using synchrotron data collected at Advanced Photon Source beamline 19-ID. After refinement with SHELX, employing anisotropic displacement parameters and riding hydrogen atoms, $R = 0.132$ and $R_{free} = 0.162$. The average coordinate estimated standard deviations are 0.021 Å and 0.038 Å for backbone atoms and side-chain atoms, respectively. Besides providing a more precise view of the α-isoform than previously available, these data permit comparison with the 0.91 Å structure determined for pike β-parvalbumin. Visualization of the anisotropic displacement parameters as thermal ellipsoids yields insight into the atomic motion within the Ca^{2+}-binding sites. The asymmetric unit includes three parvalbumin (PV) molecules. Interestingly, the EF site in one displays uncharacteristic flexibility. The ellipsoids for Asp-92 are particularly large and nonspherical, and the shape of the Ca^{2+} ellipsoid implies significant vibrational motion perpendicular to the plane defined by the four *y* and *z* ligands. The relative dearth of crystal-packing interactions in this site suggests that the heightened flexibility may be the result of diminished intermolecular contacts. The implication is that, by impeding conformational mobility, crystal-packing forces may cause serious overestimation of EF-hand rigidity. The high quality of the data permitted 11 residues to be modeled in alternative side-chain conformations, including the two core residues, Ile-97 and Leu-105. The discrete disorder observed for Ile-97 may have functional ramifications, providing a mechanism for communicating binding status between the CD and EF binding loops and between the PV metal ion-binding domain and the N-terminal AB region.

Keywords: parvalbumin; EF-hand; Ca^{2+}-binding proteins; atomic resolution crystallography

EF-hand proteins are essential components of the eukaryotic Ca^{2+}-signaling machinery (Celio et al. 1996). Certain members of this large protein family function as Ca^{2+}-dependent signal transducers, calmodulin and troponin C being the archetypal examples. Others function as soluble Ca^{2+} buffers, shaping the size and duration of the Ca^{2+} signal. The small (M_r 12,000), vertebrate-specific proteins called parvalbumins (PVs) exemplify the latter. These proteins occupy a prominent place in EF-hand protein lore—the first family members detected, purified to homogeneity, sequenced, and crystallized (Wnuk et al. 1982). Importantly, the X-ray structure of carp parvalbumin, reported by Kretsinger and Nockolds (1973), established the EF-hand structural paradigm.

Kretsinger observed that the parvalbumin molecule consists of three homologous 30-residue segments—each consisting of a central loop flanked by short, amphipathic helices (Kretsinger and Nockolds 1973). These were termed

Reprint requests to: John J. Tanner, Department of Chemistry, University of Missouri–Columbia, Columbia, MO 65211, USA; e-mail: tannerjj@missouri.edu; fax: (573) 882-2754; or Michael T. Henzl, Department of Biochemistry, University of Missouri–Columbia, Columbia, MO 65211, USA; e-mail: henzlm@missouri.edu; fax: (573) 884-4812.

Abbreviations: PV, parvalbumin; RMSD, root-mean-square difference; e.s.d., estimated standard deviation; PDB, Protein Data Bank; ADP, anisotropic displacement parameter.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.03571004.

the AB, CD, and EF domains, in reference to the flanking helical elements. Only the latter two display Ca^{2+}-binding activity; it is thought that a two-residue deletion abolished the metal ion-binding capacity of the AB domain. Within the CD and EF binding loops, the six ligands to the bound Ca^{2+} are located at the approximate vertices of an octahedron, and are indexed by Cartesian axes. Side-chain oxygen atoms furnish the $+x$, $+y$, $+z$, and $-z$ ligands. A main-chain carbonyl provides the invariant $-y$ ligand. The $-x$ position is occupied by a glutamyl side chain in the CD site and by a water molecule in the EF site. Ligation by the nearly invariant glutamyl carboxylate at $-z$ is bidentate, so that the Ca^{2+} coordination geometry is actually pentagonal bipyramidal. Kretsinger also made the paradigmatic observation that the CD and EF sites are related by a twofold symmetry axis, and are physically joined by a short segment of antiparallel β-sheet structure (Kretsinger and Nockolds 1973). With very few exceptions, the 30-residue EF-hand substructures occur in pairs, forming a so-called EF domain. The parvalbumin tertiary structure, then, includes two structural domains—the metal ion-binding domain, formed by the paired CD and EF sites, and the N-terminal AB domain. The AB domain packs tightly against the hydrophobic aspect of the CD–EF domain, preventing calmodulin-like interactions with target proteins.

In the three decades since Kretsinger and Nockolds published the structure of carp parvalbumin (Kretsinger and Nockolds 1973), over 20 additional parvalbumin structures have been deposited in the PDB. Despite our familiarity with the parvalbumin molecule, however, our understanding of parvalbumin structure–function relationships remains embryonic. For example, the rat α- and β-parvalbumin isoforms display 49% sequence identity (Epstein et al. 1986; Gillen et al. 1987), and the overall root-mean-square difference (RMSD) for their superimposed CA backbones is just 0.8–0.9 Å (based on 88 equivalent residues). Nevertheless, these two proteins exhibit markedly different divalent ion-binding properties (Hapak et al. 1989; Eberhard and Erne 1994). It is also well known that typical parvalbumin CD and EF sites, including those in rat α, behave identically in titrations with Ca^{2+} or Mg^{2+}. However, just one of the sites in α is capable of binding Na^{+} (Henzl et al. 2000), evidence that this apparent functional equality is the result of cooperativity and not a reflection of any intrinsic equivalence.

Insight into these and other issues of parvalbumin behavior will require additional high-precision physical and structural measurements. Toward that goal, Declercq et al. (1999) recently reported a 0.91 Å structure for the pike 4.10 β-isoform. The superior resolution of that structure enabled the investigators to discern multiple side-chain conformations for several internal residues, indicative of slow structural rearrangements of the hydrophobic core. They suggested that the internal mobility of the protein core is linked to the dynamics of the entire protein molecule, and that the millisecond conformational dynamics of the core could influence the kinetics of Ca^{2+} and Mg^{2+} association/dissociation.

We herein describe a 1.05 Å structure of rat α-parvalbumin. The substantially higher resolution permits significant refinement of the 2.0 Å structure presently in the PDB (McPhalen et al. 1994). Thus, the uncertainties in Ca^{2+}-ligand distances and ligand–Ca^{2+}-ligand bond angles are considerably smaller. Analysis of solvent at atomic resolution reveals 11 water molecules common to α-PV and β-PV structures. Moreover, the quality of the data allows refinement of the anisotropic displacement parameters (ADP), and visualization of the ADPs as thermal ellipsoids facilitates an evaluation of the atomic motions in the Ca^{2+}-binding sites. Finally, as observed for the pike 4.10 β-isoform, alternative side-chain conformations are also evident for several internal residues in the high-resolution structure of the rat α-isoform.

Results and Discussion

Quality of the structure and overall fold

The crystallographic asymmetric unit contains three α-PV molecules, labeled A, B, and C. With few exceptions, the protein electron density was unambiguous throughout. The entire backbone could be modeled with full occupancies in all three molecules, except for the terminal residue Ser-C109. Several surface side chains could not be modeled completely due to weak density (Lys-A44, Lys-A52, Lys-A80, Lys-B13, Lys-C36, Lys-C37, Glu-C108, and Ser-C109), and were therefore truncated. We note that Lys-13, Lys-44, and Lys-52 are truncated in the 1.54 Å shark α-PV structure (PDB code 5PAL; Roquet et al. 1992).

As expected, the protein molecules display the characteristic PV fold, which consists of six α-helices (labeled A–F) connected by loops (Fig. 1; Kretsinger and Nockolds 1973; Celio et al. 1996). The helices associate in pairs to form three domains labeled AB, CD, and EF. Calcium ions are bound in the helix-loop-helix, or EF-hand, substructures of the CD-EF domain. The two Ca^{2+}-binding sites are joined by a short segment of antiparallel β-sheet structure formed by Phe-57 and Ile-58 from the CD loop and Lys-96 and Ile-97 from the EF loop.

Several statistical indicators attest to the quality of the structure. For example, 96% of the nonglycine residues occupy the favored region of the Ramachandran plot; and there are no residues in the disallowed regions (Table 1). Most notably, the estimated standard deviations (e.s.d.'s) of atomic coordinates determined from blocked matrix refinement in SHELX are indicative of a well-refined high-resolution structure (Fig. 2). The average coordinate e.s.d. is 0.021 Å for backbone atoms, and 0.038 Å for side-chain atoms.

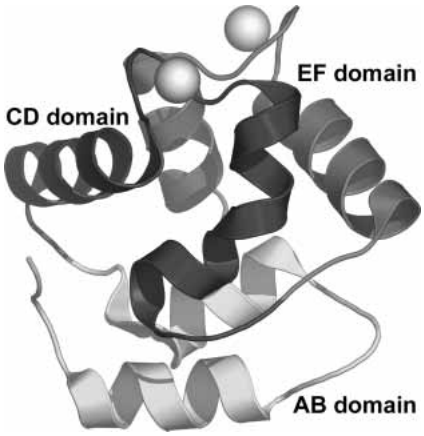

Figure 1. Ribbon drawing of one of the three protein molecules in the asymmetric unit. The bound Ca ions are indicated by spheres. Figures 1, 3, and 5 were created using PyMol (DeLano 2002).

The e.s.d. values are strongly correlated with the density of crystal contacts. Whereas regions of low e.s.d. show numerous contacts, regions of high e.s.d. display a relative dearth of contacts. For example, consider the e.s.d.'s for residues 13–37, which are much higher in the C chain than in the A and B chains (Fig. 2). Residues C13–C37 form only 12 crystal contacts, while A13–A37 and B13–B37 make 40 and 34 crystal contacts, respectively. Likewise, residues B73–B75, which have relatively high e.s.d.'s, form only a single crystal contact, while A73–A75 and C73–C75 form nine and four crystal contacts, respectively. Finally, residues 87–95, which include a portion of the EF site, display higher e.s.d.'s in the A chain (only three crystal contacts) than in the B (11 crystal contacts) and C chains (nine crystal contacts).

The protein backbone conformation is very similar to that of the 2.0 Å room temperature structure (PDB code 1RTP; McPhalen et al. 1994). For example, the RMSD between the chain A backbones of the 1.05 Å and 2.0 Å structures is only 0.34 Å (following superposition of the A chain backbone atoms). The corresponding values for the B and C chains are 0.29 Å and 0.41 Å, respectively. The backbone differences between the 1.05 Å and 2.0 Å structures within the protein core are even smaller, with RMSD values of 0.27 Å, 0.22 Å, and 0.24 Å for chains A, B, and C. The RMSDs (C_α) between our structures and the 1.54 Å shark α-PV structure (Roquet et al. 1992) are 0.7–0.8 Å; the RMSDs with the 0.91 Å pike β-PV structure (PDB code 2PVB; Declercq et al. 1999) are 0.8–0.9 Å.

The RMSDs between pairs of protein molecules in the asymmetric unit of the 1.05 Å structure are 0.6–0.7 Å, approximately two times greater than those between corresponding molecules in the 2.0 Å and 1.05 Å structures (0.3–0.4 Å). The relatively large RMSDs within the asymmetric unit are believed to reflect conformational differences resulting from disparate crystal packing environments. Indeed, this crystal form—characterized by a relatively low solvent content—contains an unusually large number of intermolecular contacts. Chains A, B, and C form 116, 119, and 118 protein–protein crystal contacts, respectively. Moreover, 45% of the noncore residues participate in at least one crystal contact. In other words, roughly half of the solvent-exposed residues engage in protein–protein crystal packing interactions.

Cryogenic treatment has evidently provoked a contraction of the unit cell (see Materials and Methods). This alteration produces an interesting difference between the 1.05 Å and 2.0 Å structures. If the asymmetric units in the two are superimposed so as to minimize the deviations in the B and C chains, the resulting RMSDs for the A, B, and C chains are 1.0 Å, 0.33 Å, and 0.43 Å. The uniquely high

Table 1. *Data collection and refinement statistics*

Wavelength (Å)	0.946
Space group	$P2_12_12_1$
Unit cell dimensions (Å)	$a = 33.8, b = 54.7, c = 153.6$
No. of protein molecules per asymmetric unit	3
Diffraction resolution (Å)	50–1.05 (1.07–1.05)[a]
No. of observations	937, 912
No. of unique reflections	132, 869
Redundancy	7.1 (3.9)[a]
Completeness (%)	99.2 (97.5)[a]
Mean I/σ_1	36.2 (4.6)[a]
R_{merge}	0.06 (0.319)[a]
No. of protein atoms	2515
No. of water molecules	373
No. of calcium ions	6
No. of sulfate ions	4
No. of ammonium ions	1
No. of PEG molecules	1
R_{cryst}	0.1325
R_{free}[b]	0.1620
RMSD from restraint target values	
Bond lengths (Å)	0.015
Angle distances (Å)	0.029
Distances from restraint planes (Å)	0.028
Zero chiral volumes (Å^2)	0.079
Nonzero chiral volumes (Å^2)	0.089
Antibumping (Å)	0.076
Rigid-bond ADP components (Å^2)	0.005
Similar ADP components (Å^2)	0.034
Approximate isotropic ADPs (Å^2)	0.064
Ramachandran plot[c]	
Favored (%)	95.6%
Allowed (%)	4.1%
Generous (%)	0.3%
Disallowed (%)	0.0%
Average isotropic B-factors (Å^2)	
Protein	11
Solvent	21

[a] Values for the outer resolution shell of data are given in parenthesis.
[b] 5% random test set (6703 reflections).
[c] Calculated using PROCHECK (Laskowski et al. 1993).

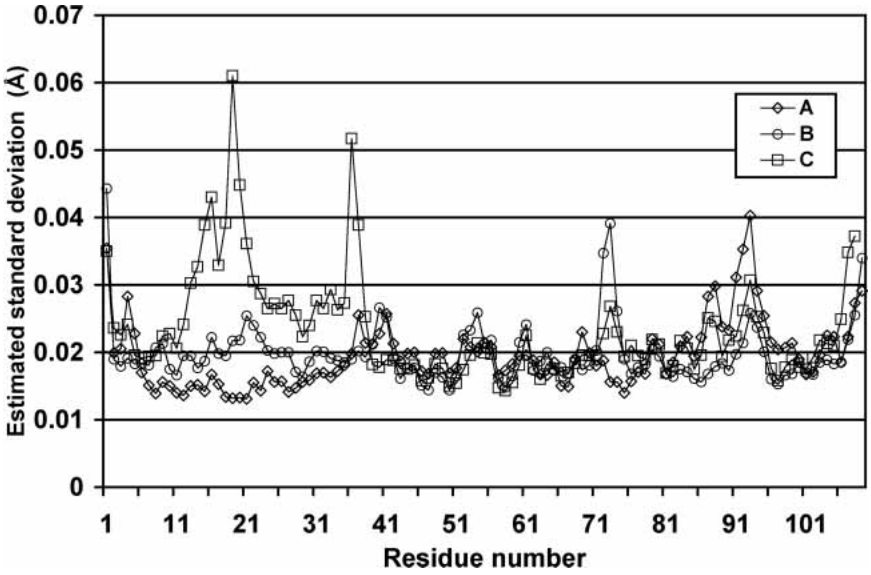

Figure 2. Estimated standard deviations (e.s.d.'s) of coordinates for backbone atoms. The e.s.d.'s were obtained from blocked matrix refinement performed in SHELX.

RMSD value for the A chain indicates a relative rigid-body movement of the A chain. This movement is significant because it breaks a crystal contact that links the EF Ca^{2+}-binding site of chain A and the backbone of chain B in the room temperature structure (discussed below).

Solvent structure: H_2O, NH_4^+, SO_4^{2-}, and PEG

The refined 1.05 Å structure includes 373 water molecules (362 at full occupancy and 11 at half occupancy). The isotropic B-factors for these water molecules are in the range 6–54 $Å^2$, with an average of 21 $Å^2$. For comparison, the 2.0 Å structure has 110 water molecules. Thus, as expected, the increased resolution affords a more detailed and complete view of protein solvation. Nine of the peaks modeled as water in the 2.0 Å structure correspond to nonaqueous solvent molecules or side chains in the 1.05 Å structure. The 1.05 Å and 2.0 Å structures have 92 water molecules in common. The average B-factor of these common water molecules in the 1.05 Å structure is 14 $Å^2$, which is lower than average B-factor of 23 $Å^2$ for the remaining water molecules in the 1.05 Å structure. The water molecules common to the two structures, then, are generally the more ordered ones, so that the increased resolution appears to provide a better picture of the less well-ordered protein-associated solvent.

Of the 373 water molecules in the asymmetric unit, 292 form at least one hydrogen bond to protein. Forty-seven water molecules form bridging hydrogen bonds between two proteins, thus directly mediating crystal packing. The low solvent content of this crystal form (38%) prompted interest in the relative influence of crystal packing on solvent structure. The crystalline environments of the three polypeptides in the unit cell differ substantially. Thus, if crystal packing considerations strongly impact solvent structure, one would anticipate a relatively small number of structurally conserved water molecules among the three chains. In fact, superposition of the A, B, and C chains with their hydration shells reveals a common set of just 29 water molecules (1.5 Å cutoff, at least one interaction in common). This finding implicates crystal packing as a major determinant of solvent structure.

A comparison of the solvent in this structure with that of the atomic resolution β-PV structure was undertaken to identify solvent sites conserved in both lineages. For this analysis, all three α-PV chains and the β-PV structure were superimposed onto a common reference structure. Conserved water molecules were strictly defined as those that (1) occupied the same location in at least one of our proteins and in the β-PV structure, and (2) had at least two common hydrogen bonding partners. The analysis revealed 11 structurally conserved water molecules and 25 conserved protein–water hydrogen bonds (Table 2). For reference, Table 2 also lists the equivalent water molecules in two other high resolution PV structures, 1.3 Å rat β-PV (PDB code 1RRO; Ahmed et al. 1993) and 1.54 Å shark α-PV (Roquet et al. 1992).

Most of the conserved hydrogen bonds involve main chain atoms and the carbonyl oxygen, in particular. Seven of the 11 conserved water molecules are likewise present in both the shark α-PV and rat β-PV structures. Although conserved solvent sites 1–7 in Table 2 have been reported earlier (Roquet et al. 1992; Declercq et al. 1999), sites 8–11 represent previously unidentified structurally conserved wa-

Table 2. *Water molecules common to four high resolution parvalbumin crystal structures*

	Water-ID								
	1RWY[a]								
Site	A	B	C	2PVB[b]	1RRO[c]	5PAL[d]	Hydrogen bond partners in 1RWY		
1	252	246	228	201	111	246	OD1-Asp-94	OD2-Asp-94	OE1-Glu-101
2	34	105	230	204	120	203	O-Leu-50	O-Glu-62	
3	57		9	203	194	236	O-Gly-89	O-Glu-101	
4	47	43	10	224		264	N-Glu-60	O-Gly-95	
5		32	33	208[e]	131	272	O-Gly-56	N-Val-99	
6	22	14	70	216	117	202	N-Ser-7	O-Val-33	OD2-Asp-10
7	35	41	18	202	115	204	O-Gly-64	N-Leu-67	O-Arg-75
8	221		196	319			O-Ser-72	O-Ala-74	
9		72		218	123	215	O-Ala-17	N-Ala-20	
10	238	51		286			O-Ile-49	NZ-Lys-54	
11	169	64	92	235			N-Ser-39	OD2-Asp-42	

[a] 1.05 Å rat α-PV.
[b] 0.91 Å pike β-PV (PDB code 2PVB; Declercq et al. 1999).
[c] 1.30 Å rat β-PV (PDB code 1RRO; Ahmed et al. 1993).
[d] 1.54 Å shark α-PV (PDB code 5PAL; Roquet et al. 1992).
[e] Symmetry mate of water 208 of the PDB file for 2PVB.

ter molecules. Site 1 has a clear functional role as the $-x$ ligand of the EF site (see Table 3). Waters 2–7 bridge residues distant in sequence, while waters 8–11 link residues relatively close in sequence.

Site 2 hydrogen bonds to the backbone of Glu-62 and Leu-50 (Table 2), which are both located in the CD Ca^{2+}-binding loop. Glu-62 provides the conserved bidentate coordination with Ca^{2+}, and Leu-50 precedes the first coordinating residue (Asp 51). Thus, site 2 joins the beginning and end of the CD loop. Site 3 forms analogous interactions in the EF domain. It hydrogen bonds to the backbone of Gly-89 and Glu-101. Thus, sites 2 and 3 are related by the pseudo-twofold symmetry of the CD-EF domain.

Sites 4 and 5 are also related by the twofold symmetry of the CD-EF domain; however, instead of being part of a single EF-hand, they bridge the CD and EF domains. These water molecules hydrogen bond to residues preceding and following the short β-sheet that joins the CD and EF loops, effectively extending the sheet in both directions and concomitantly doubling the number of hydrogen bonds between backbone atoms of the sheet. By virtue of their location, these water molecules could be involved in communication between the CD and EF sites.

Conserved solvent site 6 links residues quite distant in sequence, Ser-7, Asp-10, and Val-33. This water molecule bridges the N terminus of helix-A to the C terminus of helix-B. Note that one of its hydrogen bonding partners is the carboxylate of highly conserved Asp 10.

Sites 7 and 8 hydrogen bond to residues in the loop connecting the D and E helices. Site 9 bridges residues within the loop connecting helices A and B. Site 10 bridges the side chain of Lys-54 (a residue within the CD Ca^{2+}-binding site) and the backbone of Ile-49. Conserved water 11 helps stabilize the N terminus of helix C by bridging the first residue of the helix, Ser-39, with the $i + 3$ residue, Asp-42. Interestingly, the same bridge is formed in rat β-PV and shark α-PV, except that in those structures Gln-42 replaces both Asp-42 and the water molecule.

A solvent molecule in a crystal contact near the CD site of chain A was modeled as an ammonium ion (B-value = 10 $Å^2$), based on hydrogen bonding environment. It interacts with five protein atoms that probably serve as hydrogen bond acceptors, namely two backbone carbonyl atoms of chain C (Thr-3 and Leu-6) and three carboxylate oxygen atoms of chain A (Asp-53, Glu-59, and Asp-61). The sixth hydrogen bonding partner, a water molecule, could theoretically be either an acceptor or a donor. This site is modeled as a water molecule in the B and C chains. The structures of rat α-PV at 2.0 Å resolution (McPhalen et al. 1994) and pike 4.10 β-PV also have ammonium ions modeled at this location (Declercq et al. 1991, 1999).

The solvent structure includes four sulfate ions (E1–E4). All four are located in intermolecular contact regions, and each interacts with at least one basic residue and two water molecules. Sulfate E1 interacts with His-A26, Thr-A84, Lys-A12, and three water molecules. Sulfate E2 is stabilized by His-B48, Lys-B12, Lys-B37, and four water molecules. Sulfate E3 interacts with Lys-B54, Lys-A36, and two water molecules. Sulfate E4 is bound by Lys-B28, Ala-C79, and two water molecules.

A polyethylene glycol (PEG) fragment with molecular weight 194 Da ($HO[CH_2CH_2O]_4H$) was modeled, on the basis of a very strong tubular electron density feature in a crystal-contact region near residue Lys-C52 (Fig. 3). The

Table 3. *Ca^{2+}-binding geometry of rat α-parvalbumin determined at 1.05 Å resolution*

Ca^{2+}-ligand distances (Å)					
Position	Coordinating atom	A	B	C	Average
	CD site				
+x	Asp-51-OD1	2.26	2.28	2.27	2.27
+y	Asp-53-OD1	2.32	2.31	2.32	2.32
+z	Ser-55-OG	2.56	2.51	2.51	2.53
–y	Phe-57-O	2.33	2.33	2.34	2.33
–x	Glu-59-OE1	2.37	2.35	2.35	2.36
–z	Glu-62-OE1	2.41	2.43	2.42	2.42
–z	Glu-62-OE2	2.54	2.48	2.53	2.52
	Average	2.40	2.38	2.39	
	EF site				
+x	Asp-90-OD1	2.33	2.31	2.33	2.32
+y	Asp-92-OD1	2.34	2.38	2.34	2.35
+z	Asp-94-OD1	2.34	2.36	2.33	2.34
–y	Lys-96-O	2.37	2.35	2.36	2.36
–x	HOH-O[a]	2.46	2.37	2.38	2.40
–z	Glu-101-OE1	2.42	2.47	2.43	2.44
–z	Glu-101-OE2	2.53	2.51	2.54	2.53
	Average	2.40	2.39	2.39	

Ligand–Ca^{2+}–ligand angles (°)						
			A	B	C	Average
	CD site					
Asp-51-OD1	Ca^{2+}	Glu-59-OE1	166.0	167.0	165.4	166.1
Asp-53-OD1	Ca^{2+}	Glu-62-OE2	74.4	73.9	73.7	74.0
Ser-55-OG	Ca^{2+}	Asp-53-OD1	78.6	78.7	80.0	79.1
Phe-57-O	Ca^{2+}	Ser-55-OG	75.8	75.8	75.4	75.6
Glu-62-OE1	Ca^{2+}	Phe-57-O	79.2	81.0	80.5	80.2
Average, equatorial angles			77.0	77.3	77.4	
	EF site					
Asp-90-OD1	Ca^{2+}	HOH-O[a]	158.7	163.9	161.4	161.3
Asp-92-OD1	Ca^{2+}	Glu-101-OE2	75.6	73.6	76.1	75.1
Asp-94-OD1	Ca^{2+}	Asp-92-OD1	80.7	78.7	83.4	80.9
Lys-96-O	Ca^{2+}	Asp-94-OD1	81.0	78.6	78.8	79.5
Glu-101-OE1	Ca^{2+}	Lys-96-O	80.3	80.6	84.9	81.9
Average, equatorial angles			79.4	77.9	80.8	

Axial ligands occupy ±x positions (i.e., Asp-51 and Glu-59 in the CD site, Asp-90 and HOH in the EF site). Equatorial ligands occupy ±y and ±z positions.

[a] Water residue numbers are 252, 246, and 228 for A, B, and C, respectively.

PEG molecule adopts a horseshoe-shaped conformation, allowing it to wrap around the Lys-52 ammonium group. There are three hydrogen bonds between this ammonium group and the PEG molecule. HOH-75 and HOH-271 provide other hydrogen bonds to the PEG molecule (Fig. 3). The clarity and strength of the electron density suggested that this PEG molecule is very tightly bound. The high degree of order, along with the location between protein molecules, suggested that the PEG molecule participates in noncovalent interactions that are essential for maintenance of the crystal lattice. Indeed, this PEG molecule may be, in part, responsible for the ability of the crystal to diffract X-rays to such high resolution.

Although the electron density maps suggested the presence of a second PEG fragment near residue Lys-B27, this feature was less convincing than that in Figure 3. When attempts to model PEG into this feature were unsuccessful, an overlapping set of water molecules (i.e., 364–369) were modeled at half occupancy within this density instead.

Ca^{2+}-binding sites

The structural features of the EF-hand Ca^{2+}-binding motif have been discussed extensively (Strynadka and James 1989; McPhalen et al. 1991). The Ca^{2+} ions in both helix-loop-helix sites are heptacoordinate with pentagonal bipyramidal geometry. The Ca^{2+} ligands are listed in Table 3. The carboxyl groups of Asp and Glu account for five of the ligands in each site. Backbone carbonyl and Ser hydroxyl groups complete the coordination sphere of the CD site; the EF site features a backbone carbonyl and a water molecule.

The improved resolution of the present structure provides a more precise determination of the coordination geometry within the Ca^{2+}-binding sites than was previously available at 2.0 Å resolution. The bond distances and angles describing the Ca^{2+}-binding geometry are listed in Table 3. Although the average bond lengths and angles are similar to those of the 2.0 Å structure, there is significantly less molecule-to-molecule variation in the 1.05 Å structure. For example, in the 2.0 Å structure, the Asp-50–Ca^{2+} bond length varies from 2.08–2.35 Å (McPhalen et al. 1994). The variation in this bond length at 1.05 Å resolution is just 2.26–2.28 Å. Likewise, the variation in bond angles among the three molecules in the asymmetric unit is lower in the 1.05 Å structure than in the 2.0 Å structure. Whereas the axial bond angle in the EF site ranges from 152°–168° in the 2.0 Å structure, the variation is just 159°–164° in the present structure.

The bonds and angles of the Ca^{2+}-binding sites are virtually identical to those of the 0.91 Å β-PV structure (Declercq et al. 1999). The average bond length difference between the two structures is 0.01 Å, and the average angle

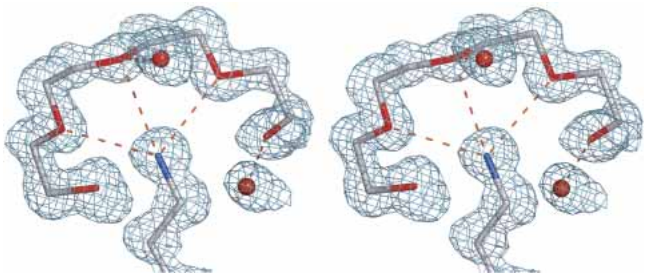

Figure 3. Stereoscopic view depicting the PEG fragment bound in a crystal contact, with 2Fo–Fc density contoured at 1σ.

difference is 1.1°. The largest bond length difference is 0.03 Å (–*x* bonds), and the largest angle difference is 1.6° (Asp94–Ca^{2+}–Asp92 angle).

Very high-resolution crystallographic data permit refinement of ADPs. The ADPs can be visualized as thermal ellipsoids, thereby providing insight into the amplitude and direction of atomic motions (Stout and Jensen 1989). For five of the six Ca^{2+}-binding sites in the asymmetric unit, the thermal ellipsoids for the Ca^{2+} ions and their oxygen ligands are generally small and nearly spherical. By contrast, the thermal ellipsoids for the atoms in the EF site of chain A are much larger than those in the other sites and decidedly nonspherical, indicating a substantially less rigid Ca^{2+}-binding site (Fig. 4). Note, in particular, that the ellipsoids for Asp-A92 are much larger and more anisotropic that those of the other Ca^{2+}-binding ligands. Also, the Ca^{2+} ion ellipsoid for the chain A EF site suggests vibration perpendicular to the equatorial plane (defined by the ligands at the ±*y* and ±*z* positions). The isotropic B-values also suggest increased flexibility in the EF site of chain A. The average isotropic B-value of the Ca^{2+} ion and its oxygen atom ligands is 13 $Å^2$ for the EF site of chain A, but only 7–8 $Å^2$ for the other five sites.

In all likelihood, the larger thermal parameters associated with this site reflect a paucity of noncovalent interactions with surrounding residues relative to the other Ca^{2+}-binding sites. Whereas the EF site of chain A (residues A90–A101) forms nine protein–protein crystal contacts, the EF sites of chains B and C form 12 and 16 crystal contacts, respectively. The CD sites of chains A, B, and C are even more highly constrained by the crystalline environment, forming 34, 13, and 26 protein–protein crystal contacts, respectively. Careful inspection of the Ca^{2+}-binding sites in the context of the crystalline environment revealed interactions that appear to be particularly important in stabilizing the EF sites of chains B and C. For example, Asp-B92 hydrogen bonds to a symmetry mate of Thr-B19, while Asp-B94 ion pairs with Lys-B96. These interactions help stabilize the EF site of the B chain, resulting in lowered thermal parameters and a more rigid binding site. This situation is even more evident in the

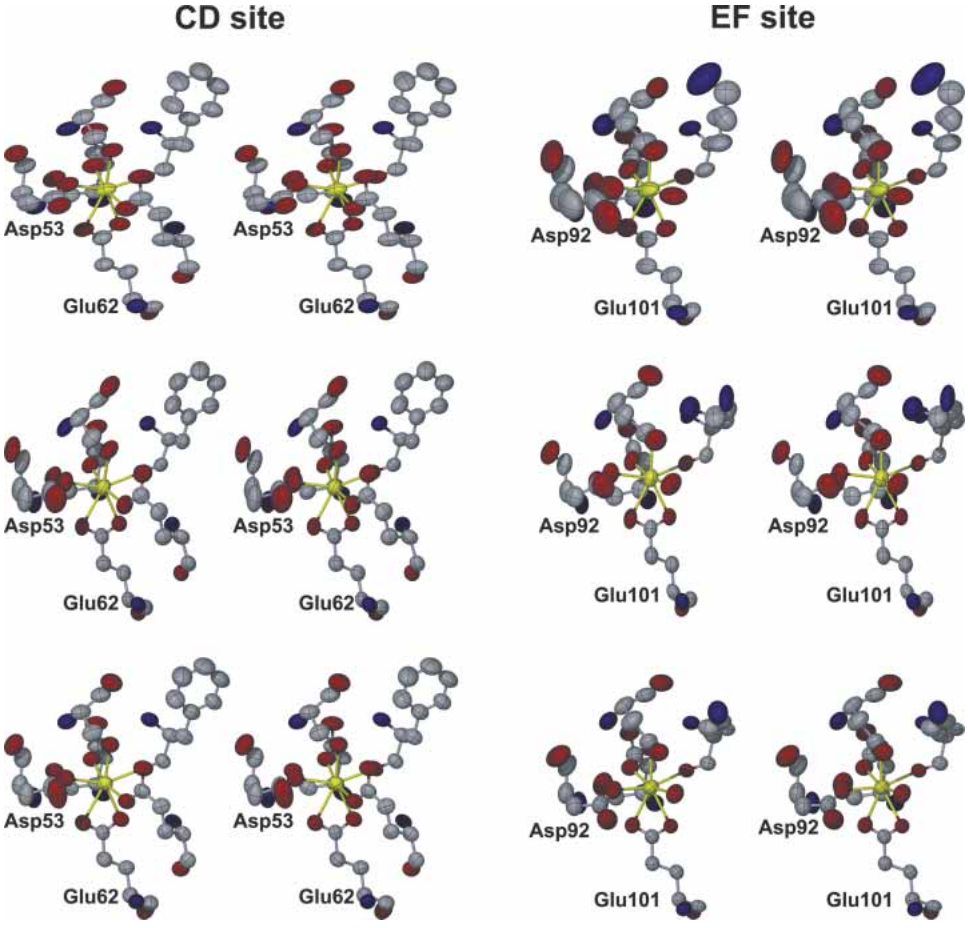

Figure 4. Stereoscopic views of the six Ca^{2+}-binding sites in the asymmetric unit, with the atoms represented as thermal ellipsoids displayed at the 50% probability level. The CD sites appear on the *left* side of the figure, and the EF sites appear on the *right* side. The *top*, *middle*, and *bottom* rows correspond to the sites of chains A, B, and C, respectively. This figure was made with X-seed (Barbour 2001).

EF site of chain C, where a PEG molecule is wedged between the Ca^{2+}-binding site and a neighboring protein molecule. The PEG molecule is within van der Waals contact (3.7–3.9 Å) of side-chain atoms of Asp-C92 and Glu-C101. Moreover, Asp-C94 hydrogen bonds to a crystallographic symmetry mate of Ser-C72, while Lys-C96 ion pairs to a symmetry mate of Asp-C73. Again, these noncovalent interactions stabilize the EF Ca^{2+}-binding site in the C chain.

These contacts are not observed in the EF site of chain A, and thus this Ca^{2+}-binding site displays higher thermal parameters. We therefore suggest (1) that the crystal structures of the EF sites in the B and C chains overestimate the rigidity of the EF site, and (2) that the intrinsic flexibility of this site is more accurately depicted by thermal ellipsoids in the A chain EF site. Furthermore, the mobility of the A chain is probably more representative of the solution behavior.

Interestingly, the A chain EF Ca^{2+}-binding site in the room temperature 2.0 Å structure does not display reduced rigidity. The average B-value for that site in the 2.0 Å structure is 8 $Å^2$, compared to 5–13 $Å^2$ for the other sites. This discrepancy between the 1.05 Å and 2.0 Å structures arises from a difference in the crystal-packing environment around the EF site of chain A. In the 2.0 Å structure, a hydrogen bond links OD2 of Asp-A94 and the backbone N—H of Ala-B79 (OD2-N distance = 2.9 Å). This interaction is absent in the 1.05 Å structure (OD2-N distance = 4.4 Å), due to the rigid-body movement of chain A caused by freezing the crystal in liquid nitrogen. Evidently, this intermolecular hydrogen bond contributes to the EF-site rigidity in the 2.0 Å structure, and its absence in the present structure allows greater flexibility.

Residues with alternative side-chain conformations

The quality of the 1.05 Å electron density map permitted the identification of 11 residues having alternative side-chain conformations (Fig. 5). The side-chain dihedral angles for these residues—all of which were modeled by single side-chain conformations in the 2.0 Å structure – are listed in Table 4. Although most of these occur on the surface of the protein, Ile-97 and Leu-105 are notable because (1) they contribute to the hydrophobic core of the protein, and (2) because the corresponding residues display two side-chain conformations in other high-resolution EF-hand protein structures. Ile-97 has been modeled with dual side-chain conformations in the 1.44 Å structure of the rat α CD-EF fragment (PVratΔ37, PDB code 1G33; Thepaut et al. 2001), the 0.91 Å β-PV structure (Declercq et al. 1999), and the 1.0 Å calmodulin structure (PDB code 1EXR, residue Ile-136; Wilson and Brünger 2000). Leu-105 displays two side-chain conformations in 2PVB (Met-105) and in 1EXR (Met-144), but only a single conformation in PVratΔ37.

It should be emphasized that alternative conformations have been modeled for Ile-97 and Leu-105 only in chain A of our structure. Although the electron density maps suggested minor secondary conformations for B97 and C97, the density was not sufficiently convincing to warrant the inclusion of two conformations. The unmistakable presence of two side-chain conformations for A97 may reflect the relatively greater flexibility of the A-chain EF site (see previous section).

The presence of buried residues with alternative side-chain conformations in β-PV has been interpreted as evidence for conformational multistates within the hydrophobic core, and it has been proposed that conversion between these multistates might represent a slow dynamical process directly related to Ca^{2+} binding and release (Declercq et al. 1999). Our results would seem to support this idea, and extend it to include α-PV as well. The calmodulin structure has a relatively large number of residues with multiple conformations (36 total). These discretely disordered residues may provide the structural plasticity that enables calmodulin

Table 4. *Dihedral angles (°) for side chains of residues having two alternative conformations (A and B)*

	Conformation A					Conformation B				
Residue	Occ.	χ_1	χ_2	χ_3	χ_4	Occ.	χ_1	χ_2	χ_3	χ_4
Lys A12	0.69	172	−171	173	164	0.31	−168	−98	−174	−179
Gln A31	0.52	−69	158	27		0.48	172	69	54	
His A48	0.61	−72	−66			0.39	171	−110		
Ser A72	0.50	59				0.50	−49			
Thr A82	0.48	63				0.52	−52			
Ile A97	0.56	−71	167			0.44	−47	−60		
Leu A105	0.46	−175	−178			0.54	−175	63		
Ile B15	0.58	−63	−76			0.42	−63	−164		
Lys B38	0.51	−64	−82	106	−172	0.49	−64	−40	−64	−168
Lys B96	0.74	60	−166	−177	−66	0.26	60	−166	−177	−170
Gln C31	0.50	−65	157	−154		0.50	−159	−91	−23	

to interact with a broad spectrum of effector molecules (Wilson and Brünger 2000).

There is presently no crystal structure available for Ca^{2+}-free rat α-PV. However, it is apparent that divalent ion binding must provoke significant conformational alterations in the molecule. Although residues 21 and 80 are well separated in the Ca^{2+}-bound protein (CA–CA distance = 9 Å), they clash in the apo-protein (Agah et al. 2003)—indicating that substantial interdomain rearrangement must accompany divalent ion binding. Furthermore, the highly cooperative divalent ion-binding behavior (Henzl et al. 2004) displayed by rat α-PV requires significant intersite rearrangement.

That Ile-97 should display alternative conformations in three distinct PV structures suggests that the dynamics of this residue might be vital to parvalbumin function. Ile-97 is ideally positioned to transmit/receive binding information—simultaneously a component of the EF Ca^{2+}-binding loop, the short β-sheet segment connecting the CD and EF sites, and the hydrophobic core. We therefore suggest that the discrete disorder of Ile-97 may provide a mechanism for transmitting binding information between the two Ca-binding sites and between the EF site and the core. Targeted site-specific mutagenesis studies could provide insight into this issue. For example, replacement of Ile-97 by Val or Ala would reduce the side-chain volume and the number of possible side-chain conformations, perhaps reducing the cooperative interaction between the two binding sites.

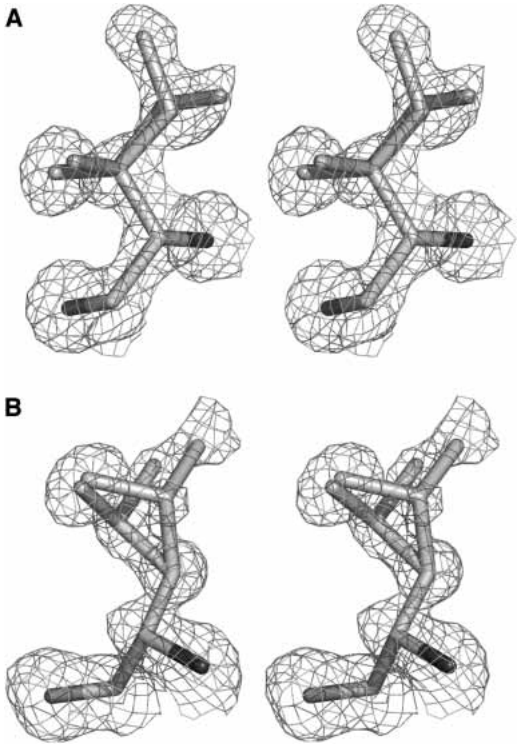

Figure 5. Stereoscopic view of core residues having dual side-chain conformations. (*A*) Ile-A97. (*B*) Leu-A105. The 2Fo–Fc map is contoured at 1σ in each figure.

Materials and methods

The purification of recombinant rat α-PV has been described previously (Henzl and Graham 1999). The homogeneity of the isolated material was assessed by SDS-PAGE and UV absorbance spectroscopy. Because rat α-PV lacks tyrosine and tryptophan, the UV spectrum is a valuable indicator of purity. Assuming an average extinction coefficient at 280 nm of 1.0 $(mg/mL)^{-1}$ cm^{-1} for contaminants, the purity of the α preparation used for crystallization exceeded 99%.

Rat α-PV was dialyzed to equilibrium against 50 mM sodium acetate, 100 μM $CaCl_2$ (pH 4.90), prior to use. Crystallization conditions were adapted from McPhalen et al. (1994). The crystals were grown at 22°C by vapor diffusion. Reservoir solutions contained $(NH_4)_2SO_4$ (80% saturation), 50 mM sodium acetate (pH 4.55–4.66), 12 mM $CaCl_2$, and PEG 600 (1%–3%). Hanging drops were formed by combining 2 μL of the protein solution (12–20 mg/mL) with an equal volume of reservoir solution. The largest crystals were produced in 3% PEG at pH 4.55. A crystal was prepared for cryogenic data collection by soaking it in a solution of 80% saturated ammonium sulfate, 11% glycerol, and 50 mM sodium acetate for a few minutes. The crystal was then picked up with a cryoloop and plunged into liquid nitrogen.

The crystals belong to the space group $P2_12_12_1$ with three molecules per asymmetric unit, as reported previously (McPhalen et al. 1994). The unit cell dimensions are a = 33.8 Å, b = 54.7 Å, c = 153.6 Å. The Matthews coefficient (Matthews 1968) is 2.0 with a solvent content of only 38%. Note that the c-axis reported here is 2.5 Å shorter than that of the room temperature cell reported previously. Freezing of the crystal presumably caused a compression of the c-axis.

A 1.55 Å data set was collected from a single crystal using an *R*-axis IV detector coupled to a Rigaku RU-H3R copper rotating-anode generator equipped with Osmic MaxFlux confocal optics and an X-stream cryogenic system. The data set consisted of 377 frames with an oscillation angle of 0.5° per frame, detector distance of 100 mm, detector θ of zero, and an exposure time of 5 min per frame. The data were processed using d*trek, as implemented in the program CrystalClear (Pflugrath 1999). The 1.55 Å data set contained 41,649 unique reflections.

A higher resolution data set was subsequently collected from the same crystal at beamline 19-ID of the Structural Biology Center at the Advanced Photon Source (APS). The data collection consisted of three scans corresponding to different values of the detector theta and goniostat κ angles. The detector distance was 120 mm, and the oscillation range was 0.5° per frame for all three scans. Scan 1 consisted of 240 frames with θ and κ set to 0, and an exposure time of 1 sec per frame. Scan 2 consisted of 240 frames with θ = 20°, κ = 0, and an exposure time of 3 sec per frame. The final scan consisted of 280 frames with θ = 20°, κ = −50°, and an exposure time of 3 sec per frame. The data were processed using HKL2000 (Otwinowski and Minor 1997). (See Table 1 for data processing statistics.) Note that the 1.05 Å data set provided a threefold increase in the number of unique reflections compared to the 1.55 Å data set.

Initial model building and refinement efforts were done with the 1.55 Å *R*-axis data set because the APS data set was not available at the time. The coordinates of the 2.0 Å structure (PDB code 1RTP, solvent removed) were input to rigid body refinement, po-

sitional refinement, and B-factor refinement performed using CNS (Brünger et al. 1998), which resulted in an *R*-factor of 0.289 for all data to 1.55 Å. No geometrical bonding restraints were imposed on the Ca^{2+} ions during refinement. The resulting phases were input to the automated building program ARP/wARP (Morris et al. 2002), which correctly built the backbone of 317 of the expected 327 residues in the asymmetric unit, with a Connectivity Index of 0.98. Side chains were added to the ARP/wARP model with guiSIDE via CCP4i (CCP4 1994). The autobuilt model was improved with several rounds of manual building in O (Jones et al. 1991) followed by refinement in CNS.

When the 1.05 Å data set became available, the best 1.55 Å model was input to refinement in SHELX-97 (Sheldrick and Schneider 1997), and the model was completed over several rounds of iterative model building and refinement in SHELX. (See Table 1 for refinement statistics.) The SHELX calculations consisted of conjugate gradient least-squares minimization, using the default effective standard deviations for geometrical restraints. The riding hydrogen and ADP features of SHELX were used, with the ADPs of solvent molecules restrained to be approximately isotropic. No geometrical bonding restraints were imposed on the Ca^{2+} ions during refinement. After the refinement had converged, a blocked-matrix least-squares refinement calculation was performed to obtain estimated standard deviations (e.s.d.'s) for atomic coordinates. For this calculation, each protein molecule was divided into overlapping blocks of 9–10 residues, with an overlap of two residues between each block. Atomic coordinate parameters and ADPs were refined for each block.

Structure analysis was performed with X-PLOR (Brünger 1992), CNS, O, Pymol (DeLano 2002), and Protein Explorer (Martz 2002). For purposes of core RMSD calculations, core residues for rat α-PV were defined previously (McPhalen et al. 1994) as those having solvent accessibility of less than 5%: 2, 6, 11, 14, 29, 30, 33–35, 46, 47, 50, 58, 63, 66, 67, 70, 74, 85, 89, 97, 102, 105, and 106. RMSD calculations were performed with CNS and the CE Web site (Shindyalov and Bourne 1998). Structural alignments utilized transformation matrices obtained from the secondary structure matching service available at the European Bioinformatics Institute Web site (http://www.ebi.ac.uk/msd-srv/ssm) (Krissinel and Henrick 2003). Analysis of protein–protein crystal contacts was done with CNS using a 3.9 Å atom-based cutoff.

PDB accession code

The atomic coordinates and structure factors have been deposited in the PDB as entry 1RWY.

Acknowledgments

This research was supported by a grant from the NSF (no. 0131166) to M.T.H. and J.J.T. We thank the Interdisciplinary Plant Group and the Genetics Area Program of the University of Missouri–Columbia for providing predoctoral support for C.A.B. We thank the personnel of APS beamline 19-ID for assistance with data collection, especially Younchang Kim. Use of the Argonne National Laboratory Structural Biology Center beamlines at the APS was supported by the U.S. Department of Energy, Office of Energy Research, under Contract No. W-31-109-ENG-38.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Agah, S., Larson, J.D., and Henzl, M.T. 2003. Impact of proline residues on parvalbumin stability. *Biochemistry* **42:** 10886–10895.

Ahmed, F.R., Rose, D.R., Evans, S.V., Pippy, M.E., and To, R. 1993. Refinement of recombinant oncomodulin at 1.30 Å resolution. *J. Mol. Biol.* **230:** 1216–1224.

Barbour, L.J. 2001. X-Seed—A software tool for supramolecular crystallography. *J. Supramol. Chem.* **1:** 189–191.

Brünger, A.T. 1992. *X-PLOR version 3.1. A system for X-ray crystallography and NMR*, p. 382. Yale University Press, New Haven, CT.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54:** 905–921.

CCP4. 1994. The CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr. D* **50:** 760–763.

Celio, M.R., Pauls, T., and Schwaller, B. 1996. *Guidebook to the calcium-binding proteins*. Oxford University Press, New York.

Declercq, J.P., Tinant, B., Parello, J., and Rambaud, J. 1991. Ionic interactions with parvalbumins. Crystal structure determination of pike 4.10 parvalbumin in four different ionic environments. *J. Mol. Biol.* **220:** 1017–1039.

Declercq, J.P., Evrard, C., Lamzin, V., and Parello, J. 1999. Crystal structure of the EF-hand parvalbumin at atomic resolution (0.91 Å) and at low temperature (100 K). Evidence for conformational multistates within the hydrophobic core. *Protein Sci.* **8:** 2194–2204.

DeLano, W.L. 2002. The PyMOL molecular graphics system. http://www.pymol.org.

Eberhard, M. and Erne, P. 1994. Calcium and magnesium binding to rat parvalbumin. *Eur. J. Biochem.* **222:** 21–26.

Epstein, P., Means, A.R., and Berchtold, M.W. 1986. Isolation of a rat parvalbumin gene and full length cDNA. *J. Biol. Chem.* **261:** 5886–5891.

Gillen, M.F., Banville, D., Rutledge, R.G., Narang, S., Seligy, V.L., Whitfield, J.F., and MacManus, J.P. 1987. A complete complementary DNA for the oncodevelopmental calcium-binding protein, oncomodulin. *J. Biol. Chem.* **262:** 5308–5312.

Hapak, R.C., Lammers, P.J., Palmisano, W.A., Birnbaum, E.R., and Henzl, M.T. 1989. Site-specific substitution of glutamate for aspartate at position 59 of rat oncomodulin. *J. Biol. Chem.* **264:** 18751–18760.

Henzl, M.T. and Graham, J.S. 1999. Conformational stabilities of the rat α- and β-parvalbumins. *FEBS Lett.* **442:** 241–245.

Henzl, M.T., Larson, J.D., and Agah, S. 2000. Influence of monovalent cations on rat α- and β-parvalbumin stabilities. *Biochemistry* **39:** 5859–5867.

———. 2004. Influence of monovalent cation identity on parvalbumin divalent ion-binding properties. *Biochemistry* **43:** 2747–2763.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47:** 110–119.

Kretsinger, R.H. and Nockolds, C.E. 1973. Carp muscle calcium-binding protein. II. Structure determination and general description. *J. Biol. Chem.* **248:** 3313–3326.

Krissinel, E. and Henrick, K. 2003. Protein structure comparison in 3D based on secondary structure matching (SSM) followed by Ca alignment, scored by a new structural similarity function. In *Proceedings of the 5th International Conference on Molecular Structural Biology* (eds. A.J. Kungl and P.J. Kungl), p. 88. Vienna.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26:** 283–291.

Martz, E. 2002. Protein explorer: Easy yet powerful macromolecular visualization. *Trends Biochem. Sci.* **27:** 107–109.

Matthews, B.W. 1968. Solvent content of protein crystals. *J. Mol. Biol.* **33:** 491–497.

McPhalen, C.A., Strynadka, N.C., and James, M.N. 1991. Calcium-binding sites in proteins: A structural perspective. *Adv. Protein Chem.* **42:** 77–144.

McPhalen, C.A., Sielecki, A.R., Santarsiero, B.D., and James, M.N. 1994. Refined crystal structure of rat parvalbumin, a mammalian α-lineage parvalbumin, at 2.0 Å resolution. *J. Mol. Biol.* **235:** 718–732.

Morris, R.J., Perrakis, A., and Lamzin, V.S. 2002. ARP/wARP's model-building algorithms. I. The main chain. *Acta Crystallogr. D* **58:** 968–975.

Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276:** 307–326.

Pflugrath, J.W. 1999. The finer things in X-ray diffraction data collection. *Acta Crystallogr. D* **55:** 1718–1725.

Roquet, F., Declercq, J.P., Tinant, B., Rambaud, J., and Parello, J. 1992. Crystal structure of the unique parvalbumin component from muscle of the leopard shark (*Triakis semifasciata*). The first X-ray study of an α-parvalbumin. *J. Mol. Biol.* **223:** 705–720.

Sheldrick, G.M. and Schneider, T.R. 1997. SHELXL: High-resolution refinement. *Methods Enzymol.* **277:** 319–343.

Shindyalov, I.N. and Bourne, P.E. 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **11:** 739–747.

Stout, G.H. and Jensen, L.H. 1989. *X-ray structure determination: A practical guide*, 2nd ed. John Wiley & Sons, New York.

Strynadka, N.C. and James, M.N. 1989. Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu. Rev. Biochem.* **58:** 951–998.

Thepaut, M., Strub, M.P., Cave, A., Baneres, J.L., Berchtold, M.W., Dumas, C., and Padilla, A. 2001. Structure of rat parvalbumin with deleted AB domain: Implications for the evolution of EF hand calcium-binding proteins and possible physiological relevance. *Proteins* **45:** 117–128.

Wilson, M.A. and Brünger, A.T. 2000. The 1.0 Å crystal structure of Ca(2+)-bound calmodulin: An analysis of disorder and implications for functionally relevant plasticity. *J. Mol. Biol.* **301:** 1237–1256.

Wnuk, W., Cox, J.A., and Stein, E.A. 1982. Parvalbumins and other soluble high-affinity calcium-binding proteins from muscle. *Calcium Cell. Funct.* **2:** 243–278.